Barış Deniz Sağlam
1558410

Term Project Proposal

Problem and task

Apple industry in the U.S. has a $15 billion annual market size. Various kinds of plant diseases cause significant economic losses. The manual diagnosis of apple plant diseases is laborious and expensive. Hence, computer vision-based methods have been developed to detect diseases from the images of leaves. The variations in imaging conditions such as backgrounds, lighting, and the large variety of visual symptoms are the main challenging aspects of these methods.

In this project, a deep learning model will be trained to detect diseases from the leaf images.

Dataset

Plant Pathology 2021-FGVC8 dataset [1], extended from Plant Pathology 2020-FGVC7 dataset [2], will be used to train the model. The dataset consists of over 18632 RGB images and labels samples, where each label may have multiple target classes of disease, since a plant may have multiple diseases at the same time. The samples are labeled by experts.
Image sizes vary from 2592 x 1728 to 5184 x 3456.

There are 5 kinds of diseases; **healthy**, **frog_eye_leaf_spot**, **powdery_mildew**, **rust**, and **scab**.

The label "complex" is used when a plant has too many diseases. The distribution of labels are as following:

scab	0.283
healthy	0.229
frog_eye_leaf_spot	0.216
complex	0.107
rust	0.103
powdery_mildew	0.063

Hence, there is imbalance among classes in the dataset.

Literature Review

Guo et al. [3] proposed a two-branch convolutional neural network with attention consistency loss which
ensures the augmentation of the images translates to the same transformation on the attention maps in the model.

This leads to superior performance in multi-label classification tasks where the augmentations may significantly change
the labels of the sample in contrast to the single-label classification task.

Chen et al. [4] use Graph Convolution Network to model label dependencies in multi-label classification tasks. Since the objects co-occur in the physical world, discovering the label dependencies significantly improves the model performance. The nodes in GCN are represented by word embeddings and the edges represent the label dependencies. The GCN builds a set classifier from this graph, which is then combined with visual features extracted with a convolutional neural network. This method may not perform well on the domain-specific labels where the word embeddings may not have rich semantic information.

Zhu et al. [5] proposed a deep neural network exploiting relations between labels by generating attention maps per label from only image-level labels. The visual features extracted by a ResNet backbone are weighted averaged by label attention maps and then mapped to the binary labels. This model performs well where there are strong correlations between locations and labels in the image such as images with common objects. For instance, the sky or clouds are mostly at the top portion of images. However, it may not achieve the same performance where there is no such spatial prior for objects in the scene.

Methods

Model architecture

Different sizes of the following model architectures will be experimented with:
- ResNet [6]
- EfficientNet [7]

Loss function

Since there are multiple labels for each sample, “Binary Cross Entropy” loss function will be used to train the model.

To deal with the imbalance in the dataset, the following techniques will be experimented with:
- Upsampling
The classes with lower frequencies will be oversampled during training so that the distribution of classes in the dataset matches the real distribution.
- Weighting losses
The losses will be weighted inversely with the frequency of the class in the dataset and in reality so that every class contributes at similar amounts to the average loss during training.

Metrics

The mean F1 score will be used as the main metric to evaluate the model's performance.

Barış Deniz Sağlam
1558410

References

1) Thapa, R., Wang, Q., Snavely, N., Belongie, S., & Khan, A. (n.d.). The Plant Pathology 2021 Challenge dataset to classify foliar disease of apples.

2) Thapa, R., Zhang, K., Snavely, N., Belongie, S., & Khan, A. (2020). The Plant Pathology Challenge 2020 data set to classify foliar disease of apples. Applications in Plant Sciences, 8(9), e11390. https://doi.org/10.1002/APS3.11390

3) Guo, H., Zheng, K., Fan, X., Yu, H., & Wang, S. (2019). Visual Attention Consistency Under Image Transforms for Multi-Label Image Classification (pp. 729–739).

4) Chen, Z.-M., Wei, X.-S., Wang, P., & Guo, Y. (2019). Multi-Label Image Recognition With Graph Convolutional Networks (pp. 5177–5186).

5) Zhu, F., Li, H., Ouyang, W., Yu, N., & Wang, X. (2017). Learning Spatial Regularization With Image-Level Supervisions for Multi-Label Image Classification (pp. 5513–5522).

6) He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition (pp. 770–778). http://image-net.org/challenges/LSVRC/2015/

7) Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks (pp. 6105–6114). PMLR. https://proceedings.mlr.press/v97/tan19a.html